

Observation of so-called “pursed-lip” and “curled-lip” utterances in Japanese, using video and MRI images

Chunyue ZHU¹, Toshiyuki SADANOBU²

¹ School of Languages and Communication, Kobe University

² Graduate School of Intercultural Studies, Kobe University

chunyuez@lion.kobe-u.ac.jp, sadanobu@kobe-u.ac.jp

Abstract

The Japanese language includes utterances described by the idioms “speaking with pursed lips” and “speaking with curled lips.” This study employs video and MRI imaging to examine the articulatory characteristics of these utterances (“utterances P” and “utterances C”, respectively) by comparing their articulation with that of “unmarked” utterances (“utterances U”). Through doing so, we arrive at the following four conclusions: (1) For the articulation of utterance P, the lips are projected outward, and rounded by expanding in the vertical direction and narrowing in the horizontal direction. (2) For the articulation of utterance C, curling the lips is not an absolute requirement. The articulation of utterance C is similar with that of utterance P in that the lips are projected outward and rounded. (3) Utterances P and C differ in two points: (a) Lips projection accompanies the lower jaw projection only in utterances P; (b) Lips in utterance P is wider than those in utterance C. (4) The shapes the lips make in utterance P, utterance C, and utterance U can be described as a circle, a horizontal rectangle, and a horizontal oval, respectively. (5) There are many facts that contradict the accepted theory that “Rounding the lips causes both lips to project outward. In reaction to this movement, the surface of the tongue is pushed toward the rear” (Koizumi 1989).

Index Terms: “pursed-lip utterances”, “curled-lip utterances”, video images, MRI images

1. Introduction

Although there has been considerable research on fundamental frequency, amplitude, and timing of speech, it was not until the 21st century that researchers began to conduct close investigation into voice-quality of speech ([1]). Previous studies on various voice qualities in Japanese everyday speech adopt either quantitative approach ([1]) or qualitative approach ([2][3]). This study is a first step toward the comprehensive description of Japanese voice qualities by using qualitative and quantitative approaches both.

The specific topic of study consists of three types of articulation in contemporary Japanese communication. These three types of articulation can be described in approximate terms as the articulation of speech described in the Japanese idiom “*kuchi o togarasete mono o iu*” (“to speak with pursed lips”), the articulation of speech described in the idiom “*kuchi o yugamete hinan suru*” (“to criticize with curled lips”), and the articulation of “unmarked” common speech. However, strictly speaking these descriptions differ a little from the actual facts of the matter.

The Japanese idiom “to speak with pursed lips” refers to a childish way of speaking that expresses displeasure or dissatisfaction ([4]). However, there is another way of speaking that is highly similar to this one, used by adults to express timidity. While the intuition of native speakers tells us

that these differ, for purposes of describing them in print they will be treated together here. This paper will attempt to make clear the features of these articulations of “speaking with pursed lips”. We will refer to these ways of speaking as “utterance P” hereinafter.

Another Japanese idiom, “to criticize with curled lips,” refers to a way of speaking that expresses a feeling of contempt. For example, the phrase “*Ore wa ee nen kedo na,’ tte, ano hito konna koto iun da yo*” (“He says, ‘I don’t care’”) spoken with curled lips expresses contempt for the object. However, the intuition of a native speaker tells us that the part spoken with curled lips is not the criticism portion (“*ano hito konna koto iun da yo*”) but the speech repeated as the object of contempt (“*Ore wa ee nen kedo na*”). This paper will attempt to make clear the features of this articulation. We will refer to utterance spoken in this way as “utterance C” hereinafter.

In contrast to utterance P and utterance C, we will identify as utterance U speech that contains no feeling of displeasure, dissatisfaction, timidity, or contempt.

This study will attempt to elucidate, through empirical methods, the actual state of the articulations of these three types of utterances. We used two high-definition video cameras positioned on the front and the side of the face to record the movements of the entire face, including the speech organs, in order to examine the external properties of the speech organs and MRI imaging to record the movements of articulation in order to examine the internal properties of the speech organs, such as the shape of the vocal tract, and we analyzed the results.

2. Methods

2.1. Experiment participant

At this stage, the experiment participant is one speaker of standard Japanese who resides in the Kansai region. This is because it was not easy to find speakers who could clear the strict requirements of MRI imaging.

2.2. Spoken text

As the spoken text, we chose the following three sentences, for which the style of speaking would be easy to recall when speaking them.

- a. Iyaa, chotto sorewa muzukashiis nee. (VIDEO, MRI)
‘That’s hard!’ (utterance P, utterance U)
- b. Orewa eenenkedo na. (VIDEO, MRI)
‘I don’t care.’ (utterance C, utterance U)
- c. Sonnano muzukashii yo. (VIDEO)
‘That’s hard!’ (utterance P, utterance C, utterance U)

2.3. Equipment used

Video camera: CANON ivis HF M32 × 2

MRI: Siemens MAGNETOM Verio (3T)

2.4. Methods of imaging

2.4.1. MRI imaging

Instead of recording video images in real time, it is possible to construct video data by repeating the same utterance 96 times, using recording that synchronizes the timing of imaging and speech using two types of trigger signals: a scan signal and a noise burst ([5][6]). In this experiment, we used this video imaging method to record images of a median sagittal plane at a resolution of $1\text{ mm} \times 1\text{ mm}$ per pixel and a thickness of 3 mm per slice, creating a video image at 60 frames per second. We recorded the speech spoken during the experiment simultaneously.

2.4.2. Video imaging

We used two high-definition video cameras with resolutions of 1920×1080 pixels, one positioned in front of the speaker and one to the side of the speaker, to record speech simultaneously.

3. Qualitative analysis of the data

3.1. Qualitative analysis of the video images

Figures 1 and 2 superimpose tracings of utterance P, utterance C, and utterance U from the images recorded using the front and side video cameras, respectively. (Solid lines: utterance U, dotted lines: utterance P, dashed lines: utterance C.)

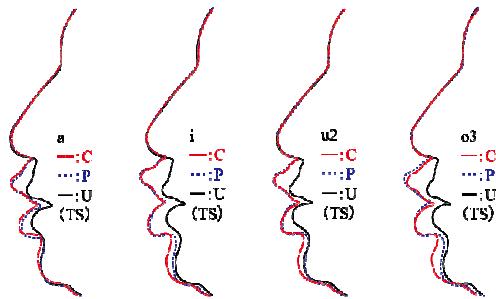


Figure 1: Lip formations (side) of utterance P, C, and U

Compared with utterance U, utterance P is characterized as prominent projection of lips, and utterance C of lower jaw.

From Figure 1, one can see the degree to which the lips project in comparison to utterance U and the changes in the position of the lower jaw when pronouncing each vowel /a/i/u/o/. Observation of vowels were conducted on their constant and steady region, and this is also the case with quantitative analysis.

We were able to confirm that while the lips project further in utterance P than in utterance U, the lower jaw projects only slightly. For this reason, the articulation of utterance P can be said to come mainly from projecting the lips.

While in utterance C the lips project to about the same degree as in utterance P, the lower jaw projects markedly forward.

Figure 2 is one example of a tracing of the outline of the lips during the steady-state parts of vowels in each utterance,

recorded from the video camera set up in front of the speaker. The shaded area in the center of each image represents the opening of the lips as seen from the front (the narrowest part of the vocal tract opening), while the second line from the outside is the boundary between the lips and the mucous membranes inside the oral cavity that can be seen when the mouth is shut at rest. The properties of the lip formation when pronouncing each of the vowel sounds /a/i/u/o/ observed from these images are outlined below.

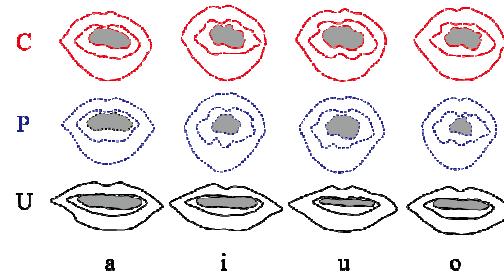


Figure 2: Lip formations (front) of utterances P, C, and U
The lips form, respectively, a circle, a horizontal rectangle, and a horizontal oval.

For utterance U, the lips cannot be described as being opened very wide in the vertical direction and the opening in the horizontal direction is unchanged from when the mouth is shut at rest. Using utterance U as the basis for articulation, in both utterance P and utterance C both the narrowness of the lips in the horizontal direction and the large opening in the vertical direction are prominent.

A look at the shape of the lip opening shows that in general the property of the lips rounding to form a circle is pronounced in the case of utterance P, while in utterance C the lips both round and extend in the horizontal direction.

In sum, for utterance P, utterance C, and utterance U the lips form, respectively, a circle, a horizontal rectangle, and a horizontal oval.

3.2. Qualitative analysis of the MRI images

Figure 3 uses MRI images to compare U and P pronunciation. It overlaps two tracings based on the part that

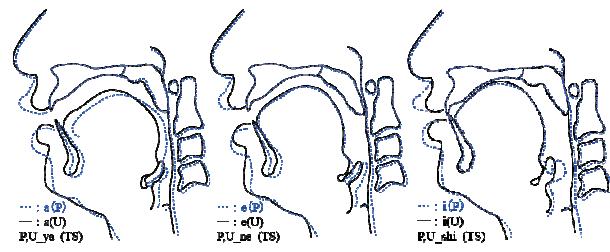


Figure 3: Form of vocal tract in utterances U and P (MRI)

Compared with utterance U and P is characterized as the forward projection of the lips. The mouth is open in the vertical direction in the left and middle drawings, but not in right drawing. In utterance P, the lower jaw also often opens. The tongue does not move toward the teeth in the middle and right drawings.

does not move during pronunciation (the front palate, including the upper jawbone).

From this figure, the following three points are clear.

(1) While the forward projection of the lips stands out more in utterance P than in utterance U, in some cases this accompanies the opening of the mouth in the vertical direction (as in the left and middle drawings) while in other cases it is due to projection alone (as in the right drawing).

(2) In utterance P, in many cases the lower jaw also opens. In such cases, the jaw-opening method basically is through a hinged movement with the temporomandibular joint serving as the axis.

(3) While the accepted theory holds that "Rounding the lips causes both lips to project outward. In reaction to this movement, the surface of the tongue is pushed toward the rear" ([7]), actually many facts contradict this accepted theory, since as seen in the middle and right pictures of Figure 3 in many cases the tongue itself does not show any movement.

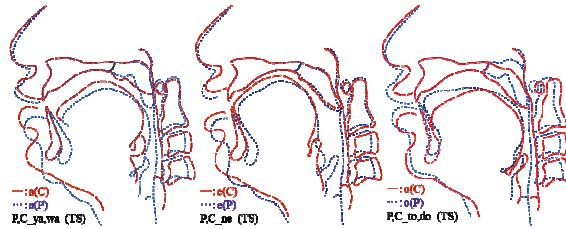


Figure 4: Shapes of the vocal tract during utterances P and C (MRI)

Utterances P and C have the projection of the lips in common.

In utterance P the lower jaw mainly simply opens through a hinged movement of the temporomandibular joint (change in angle). In utterance C the jaw lowering involves projection of the lower joint forward. In utterance C the lips often appear to thin.

Figure 4 compares P and C pronunciation by overlapping their MRI tracings in the same way as Figure 3. As we saw in Figure 1, while utterance P and utterance C have the projection of the lips in common, they differ considerably in the way the lower jaw moves. While in the former case the lower jaw mainly simply opens through a hinged movement of the temporomandibular joint (change in angle), the latter case involves projection of the lower joint forward. In addition, in many cases in utterance C the lips appear to thin. This probably is because in the case of utterance C the lips not only project forward but also elongate in the horizontal direction. (Also see Figure 2.) Tongue back of vowels /a/ and /e/ is raised much higher in utterance C than in utterance P.

4. Quantitative analysis of the data

We measured the angles and vertical and horizontal widths of the lips in the video images taken from the front and side views and compared and analyzed the findings.

4.1. Measurement method

In our quantitative analysis of the video data filmed from the front, we measured and recorded the figures for the following five points.

(1) Horizontal width of the outer perimeter of the lips (AB

in Figure 5)

(2) Height of the outer perimeter of the lips (CD in Figure 5)

(3) Horizontal width of the lip opening (narrowest part) (EF in Figure 5)

(4) Height of the lip opening (narrowest part) (GH in Figure 5)

(5) Horizontal angle of the lips (JK in Figure 5)

We also measured the following point from the video filmed from the side.

(6) Projection of lips from the face surface (MN in Figure 5)

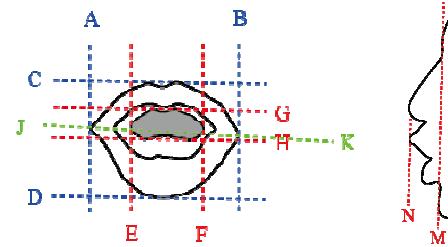


Figure 5: Measurement of each part of the lips

4.2. Measurement results and analysis

Selecting still images (144 in total) of the four vowel sounds /a//i//u//o/ from the video recorded from two cameras, one in front and one on the side, of the phrase *sonnano muzukashii yo* ("that's hard!") spoken six times each as utterance P, utterance C, and utterance U, we measured the six items listed under 4.1 above and recorded the resulting values. We took the average values for each of the four vowels in the three groups utterance P, utterance C, and utterance U and studied each parameter as described below.

4.2.1. Projection of lips

Identifying a line connecting the two points of the forehead and the jaw (M in Figure 5) to be the surface of the face, we measured the shortest distance from that line to the tip of the projected lips (N in Figure 5) as the projection of the lips. As shown in Figure 6, for each vowel sound the ranking from highest to lowest, of the extent of projection of the lips was as follows: utterance P > utterance C > utterance U. Since as we saw in Figures 1 and 4 the jaw moves forward in a horizontal direction in utterance C, its figure for pure projection is slightly lower than that for utterance P.

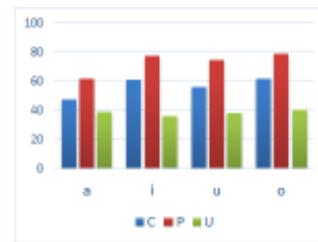


Figure 6: Projection of lips

The extent of projection of the lips was:
utterance P > utterance C > utterance U.

4.2.2. Lip height and horizontal width

Figure 7 shows the ratio of the width of the lips' outer perimeter ((1) under 4.1) to the height ((2) under 4.1), while Figure 8 shows the ratio of the width of the lip opening at its narrowest spot ((3) under 4.1) to its height ((4) under 4.1). Since the scale on the horizontal axis shows the size of the figure for the horizontal width as a multiple of the height, the higher this number the more the lips will have a wide shape, and likewise the closer this number is to 1 the more the lips will have a rounded shape.

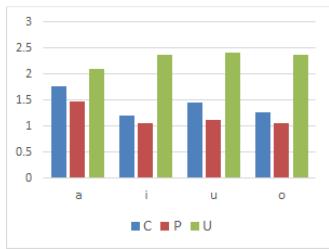


Figure 7: W/H ratio of lip outer perimeter

Utterance U shows the flattest shape and utterance P the closest to a circle. Although the shape for utterance C is close to that of a circle, the lips are wider than they are for utterance P.

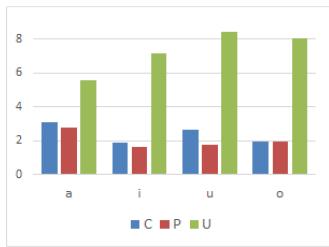


Figure 8: W/H ratio of lip opening

As well as Figure 7, utterance U shows the flattest shape and utterance P the closest to a circle. Although Utterance C is close to utterance P, the lips are wider than they are for utterance P.

As shown in Figures 7 and 8, it is clear that to one degree or another the vowels all show similar tendencies. That is, the shapes the lips form are similar in both outer perimeter and opening, with utterance U showing the flattest shape and utterance P the closest to a circle. Although the shape for utterance C is close to that of a circle, the lips are wider than they are for utterance P.

4.2.3. Lip angle

Figure 9 shows the average angle of the lips ((5) under 4.1) for each utterance. The scale on the horizontal axis is degree of inclination, and the fact that this value is negative for all utterances in each group means that the lips are inclined downward by 3° or more for each utterance.

As seen in Figure 9, in a so-called curled-lip utterance (utterance C), the lips actually are more curled than the others only for the vowel /a/, while for each of the vowels /i//u//o/ the lips are more curled in a pursed-lip utterance (utterance P). This suggests that curling the angle of the mouth is not an

absolute requirement of utterance C.

5. Conclusions

From the studies described under Section 3 and Section 4, we reached the following five conclusions.

(1) For the articulation of utterance P, the lips are projected outward, and rounded by expanding in the vertical direction and narrowing in the horizontal direction.

(2) For the articulation of utterance C, curling the lips is not an absolute requirement. The articulation of utterance C is similar with that of utterance P in that the lips are projected outward and rounded.

(3) Utterances P and C differ in two points: (a) Lips projection accompanies the lower jaw projection only in utterances P; (b) Lips in utterance P is wider than those in utterance C.

(4) The shapes the lips make in utterance P, utterance C, and utterance U can be described as a circle, a horizontal rectangle, and a horizontal oval, respectively.

(5) There are many facts that contradict the accepted theory that "Rounding the lips causes both lips to project outward. In reaction to this movement, the surface of the tongue is pushed toward the rear" ([7]).

6. Future considerations

While there are two types of utterance P —child and adult — in natural conversation this study has not taken into consideration the differences between the two. Although we have recorded video and MRI images of all four utterances of utterance P (children), utterance P (adult), utterance C, and utterance U, the measurement process for these recordings still is underway and could not be completed in time for this paper. We would like to address the results at the next opportunity.

7. Acknowledgements

We thank our "Onsei Bunpou Kenkyukai" colleagues, especially late Miyoko Sugito for extensive advice and for helping in many other ways.

This work was partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Scientific Research (A), 23320087.



Figure 9: Lip angle

The lips of utterance C are more curled than the others only for the vowel /a/, while for each of the vowels /i//u//o/ the lips are more curled in utterance P.

8. References

- [1] Campbell, N., and Mokhtari, P., "Voice quality: the 4th prosodic dimension", in Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS'03), Barcelona, Spain, 2417-2420, 2003.
- [2] Sadanobu, T., "A natural history of Japanese pressed voice", *Journal of the Phonetic Society of Japan*, 8-1: 29-44, 2004.
- [3] Sadanobu, T. *Sasayaku Koibito, Rikimu Repootaa: Kuchi nonakano Bunka*. Tokyo: Iwanami, 2005.
- [4] *Jitsuyou Nihongo Hyougen Jiten*. <http://www.practical-japanese.com/>
- [5] Masaki, S., Tiede, M. Honda, K., Shimada, Y., Fujimoto, I., Nakamura, Y., and Ninomiya, N., "MRI-based speech production study using a synchronized sampling method", *Journal of the Acoustic Society of Japan. (E)*, 20: 375-379, 1999.
- [6] Honda, K., "MRI niyoru hatsuwakikan no kansoku", Onsei Bunpou Kenkyukai (ed.), *Bunpou to Onsei*, 5, Tokyo: Kurosoi, 47-58, 2006.
- [7] Koizumi, T., "Onsei to on'in", Sugito M. (ed.), *Kouza Nihongo to Nihongo Kyouiku 2: Nihongo no Onsei, On'in*, Book 1, Tokyo: Meiji Shoin, 1-20, 1989.